APPLICATION FOR

UNITED STATES LETTERS PATENT

for

METHOD AND APPARATUS FOR ENCODING CONTROL DATA IN A VIDEO DATA STREAM

Inventor

Jeffrey A. Schriebman

RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application Serial No. 60/124,696 filed March 15, 1999.

FIELD OF THE INVENTION

The present invention relates to closed-captioning and other data appearing on the vertical blanking interval of a video image signal. More particularly, the present invention relates to an apparatus and method for the addition of closed captioning and other control data to a video or television program.

BACKGROUND OF THE INVENTION

Closed-captioning is the synchronized display of textual data with the underlying video program. In other words, closed-captioning allows a program viewer to see the video image on screen, while simultaneously reading the underlying dialogue instead of having to hear it. Traditional techniques for adding closed-captions to video data require expensive, special-purpose synchronization software and encoder hardware.

The closed-captioning of video programs is typically done in two stages. In the first stage, an enhanced dialog text script is created in a word processor and an association is made between each piece of text and a specific time sequence within the video program during which the text is to be displayed. This synchronization process is typically done by a skilled technician using a computer and associated software to associate a time code to the desired text.

In the second stage, the time-coded closed-caption data is merged with the underlying video program. Prior art methods of merging the closed-captioning and video

image data require the use of dedicated and expensive closed-captioning systems. See Figure 1. These systems include separate hardware, typically referred to as a closed-caption encoder 3. In a typical prior art system, the encoder 3 accepts a video signal from a playback VCR 2 playing a master video tape and simultaneously accepts from a computer 1 the closed captioning instructions created in the first stage. When the encoder 3 recognizes the appropriate time code it formats the associated textual/closed-caption data according to the Electronic Industries Association Standard EIA-608, Recommended Practice for Line 21 Data Service (incorporated by reference herein), and superimposes the formatted data onto the video data stream emanating from the master tape. A record VCR 4 records the output of the encoder. In this manner, the closed-caption data is placed onto line 21 of Vertical Blanking Interval (VBI) of the video data signal by the encoder. The resulting output video tape is designated as a close-captioned encoded submaster video tape.

The U.S. Federal Communications Commission specified through the Television Decoder Circuitry Act of 1990 (Decoder Act) and subsequent FCC legislation that all television receivers imported into the United States or sold in interstate commerce after July 1, 1993 with a screen size greater than 13 inches must be capable of decoding and displaying closed-captions. The Decoder Act and corresponding regulations were put into effect to make closed-captions more readily available to the hearing impaired and to provide the television broadcast industry with an ever expanding base of closed-caption capable television receivers.

From a slow beginning closed-captioned programming has gradually gained momentum to a point where television programming that contains closed-captioning is the norm. Prior to the development of the present invention, a video production house desiring to add closed-captioning to its program had three options. First, the production house could send a master tape of its program with the dialog script to a third-party closed-captioning company. Second, the production house could purchase closed-captioning software and generate the commands required to drive a hardware encoder that is operated by a third-party closed-captioning service. Lastly, the production house could purchase expensive closed captioning equipment and perform the entire closed-captioning process itself.

Faced with this type of up front capital expenditure and the costs of the various

closed-captioning options, it is desirable to find an alternate way of close-captioning a video program. Furthermore, since the use of the vertical blanking interval for the transmission of other control data, such as Content Advisory information and Uniform Resource Locator information in Interactive Television programs, is becoming more widespread, an effective and inexpensive method and apparatus are needed to add such control data to the underlying video image data.

SUMMARY OF THE INVENTION

Video programming is increasingly being edited on digital computers. Computer based video non-linear editing (NLE) systems can store in digital form an entire video program. Some digital video editing systems typically store only the image portion of the video program and regenerate the vertical blanking interval (VBI) when displaying video material. However, some non-linear editing systems digitize the last line of the vertical blanking interval and allow storage, editing and playback of that line. For example, certain industry standard digital video editing systems such as the Apple Macintosh based Avid Media Composer series with the Advanced Broadcast Video Board (ABVB) manufactured by Avid Technology, Inc. and the Apple Macintosh based Media 100 series NLE with software Release 5 manufactured by Media 100, Inc. will map a portion of the visual image area of the editing system into the VBI. This mapping occurs near the first line of the visual image so that pictorial data properly positioned near line 1 of the visual image within the editing system can be mapped to the odd and even lines of Line 21 of the VBI of a television signal. The present invention, in one embodiment, takes advantage of the above-mentioned capabilities of these and other non-linear editing systems.

The present invention provides methods and apparatuses for inserting closed-caption and/or other control data into video stream data. One embodiment inserts control data into the vertical blanking interval of a video image data stream without the use of special encoding hardware. One embodiment of the present invention provides a software solution to inserting VBI data service codes onto video programming using non-linear editing system hardware, which many video production houses already possess. One method of the present invention generally comprises generating an image data file including a pixel representation of the desired control data and merging the control image data file with an

underlying program video image data file. Control data is any non-video information normally appearing in the vertical blanking interval, or other non-image portion, of a television signal that is processed by a decoder or controller, including, but not limited to, closed captioning information, Content Advisory information, and Uniform Resource Locator (URL) information.

The apparatus, in one embodiment of the present invention, facilitates input of closed-caption or other control data from a WYSIWYG text editor, other dialog boxes, or external data files containing control data and command directives. Other control data may include, for example, Internet Uniform Resource Locator information used by an Interactive Television controller and Content Advisory information, as well as program name, program type, program length and time-in-show information. This information can be manually entered using the text editor or an on-screen window, or accepted from a preexisting computer text file containing command directives that instruct how and when the data is to appear in the vertical blanking interval of the video signal. From this information, the present invention generates a data structure that temporally arranges the control data as appropriate for further use. In one embodiment, the apparatus creates an image data file which encodes the closed caption information, Content Advisory information and/or Interactive Television Link information. In one embodiment, the present invention generates an image data file comprising a series of visual gray scale representations of the binary codes that represent the Line 21 waveform of the VBI. Of course, the present invention can be applied to other waveform protocols and to other lines of the VBI. According to the present invention, a non-linear video editor of the type described above recognizes the resulting control image data file as standard digital video picture data. Accordingly and in one embodiment, the user merges the control image data file and the video program using standard operations provided by the non-linear video editing system.

In some embodiments, the present invention encodes closed caption information according to the format disclosed in Electronic Industries Association Standard EIA-608, Recommended Practice for Line 21 Data Service. Preferred embodiments of the present invention encode program rating information according to the specifications set forth in the Transport of Content Advisory Information using Extended Data Service (XDS) EIA-744-A,

which is incorporated by reference herein. Lastly, certain embodiments of the present invention encode interactive television link information according to the Transport of Internet Uniform Resource Locator (URL) Information Using Text-2 (T-2) Service EIA-746-A, which is incorporated by reference herein.

DETAILED DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram of a typical prior art closed captioning system.

Figure 2 illustrates the Line 21 Waveform according to EIA-608 specifications.

Figures 3A, 3B and 3C comprise a flow chart diagram of a process for receiving and placing control data into two data arrays corresponding to Field 1 and Field 2 of Line 21 of the vertical blanking interval.

Figure 4 is a flow chart diagram illustrating a method for generating pictorial representations of the Field 1 and Field 2 packet waveforms of Line 21 using the data contained in the arrays which contain control data.

Figure 5 shows a dialog box called "Control Window" illustrating the various control data input options provided by one embodiment of the present invention.

Figure 6 illustrates a caption window box, which according to one embodiment of the present invention, allows for the entry of one closed caption.

Figure 7 shows a text window box into which command directives have been entered.

Figure 8 shows a dialog box presented by an apparatus according to the present invention which allows for the entry of Television Parental Guideline information.

Figure 9 is a dialog box allowing for entry of Motion Picture Association Guideline ratings.

Figure 10 shows a dialog box providing for the entry of ratings according to the Canadian English Language Rating System.

Figure 11 shows a dialog box providing for the entry of ratings according to the Canadian French Language Rating System.

Figure 12 shows a dialog box that facilitates entry of Program Information.

Figure 13 illustrates a dialog box allowing for entry of Uniform Resource Locator information.

Figure 14 illustrates a dialog box which allows the user to specify the name and address of the output file, as well as whether the resulting picture images containing the gray-scale waveforms will be in a QuickTime movie or PICT file format.

Figure 15 shows a dialog box in which the user sets the lines in the picture images on which the gray scale waveforms will be placed.

Figure 16 illustrates an apparatus embodying the principles of the present invention.

Figure 17 is a flow chart diagram setting forth a method for adding control data to a digital video data stream.

DETAILED DESCRIPTION OF THE INVENTION

In the art, data service information, including closed-captions, content advisory information, and Internet Uniform Resource Locator (URL) information, is transmitted simultaneously with the video image itself on the even and odd video fields of Line 21 of the vertical blanking interval of a television signal. Embodiments of the present invention insert control data into data structures whose elements represent points in time or image frames. The data structure is used, in one embodiment, to create a gray-scale, pictorial representation of the Field 1 and Field 2 Line 21 waveforms that include the desired control data. (*See* Figure 2). The Field 1 and Field 2 gray scale data waveforms are placed into a digital picture image. This allows a non-linear editing system to map the gray-scale waveform onto Line 21 of the vertical blanking interval of the underlying video image data to add the desired control data to the video output. The present invention, however, is not limited to encoding data to Line 21 of the vertical blanking interval and can be used to encode control data on any non-video portion of a video signal. In another embodiment, the control data as temporally arranged in the data structure is used to encode a digital video data stream with closed captioning and other control data.

Figure 16 illustrates an embodiment of the apparatus according to the present invention. As Figure 16 shows, one embodiment comprises video playback and recording device 40 operably connected to computer 20. In the embodiment of Figure 16, computer 20 includes control data input module 22 and scripting module 24. In embodiments used in connection with non-linear editing systems recording analog video streams, computer 20 further includes data file waveform generator 26. In one embodiment, computer 20

includes digital video editing module 30. In another embodiment, digital video editing module 30 can be a device physically separate from, but operably connected to, computer 20. In the embodiment shown, the apparatus also includes keyboard 60 and display device 50 and video monitor 55. Other embodiments may further include other input devices such as a mouse or joystick controlling a pointer displayed on display device 50. In one embodiment, the functionality of the present invention can be incorporated into conventional video editing systems.

Computer 20 can be any computer, special-purpose computing device, or any other suitable device for performing the required functionality. In one embodiment, user computer 20 includes at least one processor, a data storage system (including volatile and non-volatile memory, such as a hard drive), at least one input/output device, such as a floppy disk-drive and/or CD-ROM. In one embodiment, video playback and recording device 40 is a VCR device. However, video playback and recording device 40 can be any suitable device for playing back and recording video stream data or digital video stream data stored on any suitable medium. In addition, the functionality performed by computer 20 may be implemented in hardware or software, or a combination of both. In one embodiment, control data input module 22, scripting module 24, data file waveform generator 26, and digital video editing module 30 comprise executable program code, which when loaded RAM are executed by computer 20.

A. Inputting Control Data

1. User Interface for Inputting Control Data

A first aspect of the present invention is the acquisition of desired control data and its arrangement in a data structure. In one embodiment, control data input module 22 includes functionality providing user interfaces facilitating the input of control data. Figure 5 illustrates a dialog box presented by an embodiment of the present invention which allows the user to choose the form in which control data is inputted. Certain embodiments of the present invention allow for three distinct sources of input: (a) a WYSIWYG editor that allows the user to enter, position and modify one closed-caption (Figure 6); (b) an on-screen portion that allows the user to enter control data and command directives into a computer Window (Figure 7); or (c) an external data file that can contain a list of previously prepared

control data script comprising control data and command directives (see "File" in Figure 5). The external data file can be an ASCII text file or a closed-caption description file prepared by another closed-caption vendor.

In one embodiment, the WYSIWYG editor provides for input, screen positioning and editing of closed-caption text. In one preferred embodiment, this is done on a 15 column by 32 row screen matrix. Figure 6 illustrates a caption window where each box represents two character spaces and each half-box represents one character space. Closed-caption characters may have specific colors, or specific attributes such as italic, underline or flash. In one embodiment, attribute information associated with each character is stored together with the character. For this reason, the internal storage of each character in preferred forms of the WYSIWYG editor consists of a 16 bit quantity; 8 bits for the character and 8 bits for its attribute. In addition, as Figure 5 shows, the user may select standard closed-caption display options for use with the closed-caption text to be displayed. More specifically, the user may select "pop-up" or "paint-on" display modes for the closed-caption text, as well as "roll-up 2," "roll-up 3" and "roll-up 4" display modes. In one embodiment, the selection of a particular display mode affects the command and special directives that control entry of control data into the output arrays and, ultimately, how the closed-captioning is displayed on the television screen. (*See* discussion, below.)

As Figure 5 indicates, embodiments of the present invention also include the ability to generate ratings for the Television Parental Guidelines system, the Motion Picture Association of America (MPAA) Guidelines system, the Canadian English Language Rating System, and the Canadian French Language Rating System. Certain embodiments feature dialog boxes (*see* Figures 8, 9, 10 and 11) that allow the user to enter the desired rating information. Other embodiments include dialog boxes that allow the user to generate interactive television links or URLs. (*See* Figure 13.)

In addition, the control data described above can be manually scripted using an on-screen window (Figure 7) or can be contained in a previously prepared external data file containing the closed-caption, Content Advisory information and any other control data. When the on-screen window or an external data file is used, one embodiment parses the command directives to enter the associated control data in the appropriate elements in two

output arrays, described below. In one form, command directives consist of instructions and time code specifications that describe how and when the control data information is to be carried on Line 21 of the VBI. These command directives or instructions allow for the selective and/or appropriate placement in time in the video data stream closed-caption data, content advisory information (*e.g.*, MMPA program rating specification), Program Title, Type and Time-In-Show, and Internet Uniform Resource Locator (URL) information (*i.e.*, Interactive Television Link specifications). In one embodiment, these command directives are a language used to instruct the temporal arrangement and storage of control data for later use. In one embodiment, the present invention accepts control data scripts including control data and command directives from third-party systems. In order to interpret and use previously entered closed-caption data using a third-party system, embodiments of the present invention include functionality for parsing third-party scripts comprising control data and command directives to complete the output arrays that are used to create the control data waveforms.

In addition, as Figure 5 indicates, the apparatus of the present invention allows for certain options when the user manually scripts the entry of control data. The "Delay Value" option allows the user to omit the entry of time codes. In this instance, if no time code is specified with a particular command directive or control data, the apparatus assigns a time code for a particular caption according to the delay specified in the Delay Value box. For example, if pop-up or paint-on closed caption displays are used, the Delay Value is the time between successive pop-up or paint-on captions. The apparatus loads each closed-caption into the data arrays, discussed below, according to the delay value, if no time code is specified. Furthermore, if roll-up display is used, the Delay Value represents the time between the display of each line in the roll-up caption.

In addition, the Duration value sets the minimum running time of the output. More specifically, the Duration value allows the specification of an MPAA rating, for example, only once and its placement into the time-points of the data arrays spanning the specified duration. Furthermore, the "Time Code Offset" option allows for the re-mapping of time code values that are beyond the one-hour value into values which the apparatus can use.

2. Data Structures for Storing Control Data

As discussed above, the present invention inserts the inputted control data into a data structure contained in the memory of computer 20. In one embodiment, scripting module 24 parses a control data script containing control data and command directives and inserts control data into a data structure as specified by the associated command directives. In one form, one or more elements of the data structure correspond to a time point or video frame in the underlying video program data.

In one embodiment for use with EIA-608 compliant devices, the data structure comprises two linear arrays of sufficient capacity to store one hour of the Field 1 packet data values and one hour of the Field 2 packet data values, respectively. Of course, any suitable data structure can be used. In NTSC non-drop frame video thirty individual video frames are assigned to each second. Therefore, one hour of video comprises 108,000 (30 x 60 x 60) frames. In drop frame nomenclature this number is slightly less. The EIA-608 standard specifies that two 8-bit characters are to be sent with each of the two field packets of each video frame. The memory required to send Line 21 data for both fields for one hour's worth of video frames is, therefore, (2 x 2 x 108,000) 432,000 bytes. Accordingly, one embodiment of the present invention includes two arrays: one for Field 1 packet data and one for Field 2 packet data. Each pair of data elements in each array corresponds with a particular time code beginning with 00:00:00:00 and ending with 1:00:00:00 (Hours:Minutes:Seconds:Frames).

Figures 3A, 3B, and 3C set forth a method for inserting control data into a data structure of the present invention. In one embodiment, scripting module 24 initializes the odd and even arrays and initializes the odd and even array pointers (Figure 3A, step 100). In one embodiment, the two linear data arrays are filled as follows: If closed-caption input is taken from the WYSIWIG editor (step 102), the appropriate closed-caption data values are generated according to EIA-608 specifications and the FCC Report and Order of April 12, 1991 (FCC 91-119), setting forth the standard control codes for closed-caption decoders, (which is incorporated herein by reference) and placed sequentially into the Field 1 packet data array starting at 0:00:00:00 (step 104). If Content Advisory (Figure 3C, step 138), interactive television links and/or program information (step 142) is specified using the dialog boxes in Figure 5, 8, 9, 10, 11, 12 and 13, data values according to the respective

specifications are generated and placed sequentially into the appropriate data array (steps 140 and 144, respectively). The conversion of control data into the appropriate control codes can occur either when the control data is inserted into the data structure or when pixel representations of the control data are generated. In a preferred embodiment, however, conversion of control data into the appropriate control codes is performed when the control data is inserted into the data structure.

If input is taken from a computer window or from an external file (Figure 3A, step 106), the command directives are parsed and control data is placed into the appropriate Field 1 or Field 2 packet data array at the appropriate time-point (*see* steps 108-136). The EIA-608 specification, incorporated herein by reference, specifies in which Field certain control data should be placed for embodiments complying with such specifications. In one embodiment, captions and Interactive Television Links are placed first due to their critical timing nature. Next, program rating information and program information are placed into the arrays. In addition and in one embodiment, a user may use the dialog boxes to fill the arrays with Content Advisory or other information and then manually script different Content Advisory or other information to be entered at specific time points and over-write the previously entered data in the control window.

B. Generating Pixel Representations of Control Data

In the embodiment described above, each pair of characters in the respective output arrays (Field 1 and Field 2 packet data) represents the data for one frame at a particular time point that will appear on Line 21 of the VBI. Once all of the data is encoded and inserted into the output arrays, one embodiment of the present invention generates pictorial representations in gray-scale of the Line 21 waveforms for each frame. In one embodiment, the Field 1 and Field 2 packet gray scale data waveforms are placed into a digital picture image that is 720 pixels wide by 486 pixels high. The dimensions of the image frame generated by the invention are not limited by the invention; rather, the dimensions of the image frames depend on the requirements of the media and video editing system. These frames are then merged as part of a multimedia file, such as a QuickTime® file, or stored separately as sequentially-numbered image files, such as PICT files. Of course, any suitable digital image, movie or multimedia file format, such as the AVI file format, can be used in

the present invention.

Figure 4 illustrates an embodiment of a method for generating pixel representations of the control data using the data structure described above. In one embodiment, the user is prompted for the form of the output (Figure 4, step 202). In one embodiment, the user can specify the output as a QuickTime multimedia file or as a series of sequentially numbered image files, such as a PICT file. Next, the pointers for the Field 1 and Field 2 packet arrays are initialized (step 204). If the pointer is not at the end of the Field 1 and Field 2 packet arrays (step 206), data file waveform generator 26 retrieves the control data from the Field 1 and/or Field 2 packet arrays (step 208) and constructs the Line 21 waveforms for that frame (step 214). Specifically, the binary representation of the two control data values from the Field 1 array are encoded into a linear series of pixels. In one embodiment, control data is encoded into a linear series of 720 pixels according to EIA-608 specification, including the Clock Run-In and Start Bits. The resulting pixel representations are, in one embodiment, copied onto at least one line of the output image frame. Next, the binary representation of the two character data values from the Field 2 array are encoded in a similar manner and put onto at least one line of the output image frame. The resulting image file is then converted and output as a QuickTime frame or as a sequentially numbered file in PICT image format (step 216). Of course, any suitable file format, such as the AVI file format, can be used. This process is repeated until the pointer is at the end of the data arrays. (*See* Figure 4, steps 220, 206, and 222.)

In one embodiment, when the user specifies a multimedia output (step 202), such as a QuickTime file or other file format that supports references to previous frames, data file waveform generator 26 determines whether the control data for the current frame has been used in a previous frame (step 210). If so, data file waveform generator 26 builds a reference to the previous video frame (step 212).

To be used by currently available television systems, the Field 1 and Field 2 packet waveforms must adhere to the standards set forth in Reference EIA-608. *See* Figure 2. Starting with the Clock Run-In Bits and Start Bits, a gray scale representation of the waveform shown in Figure 2 is constructed. Typically, some non-linear editing systems strip off the H Sync and Color Burst portions of the Line 21 waveform. Accordingly, those

portions of the Line 21 waveform need not be recreated in the gray-scale waveform in embodiments that are used in conjunction with such NLEs. As more fully described below, gray-scale bit representations of the control data stored in the two output arrays are placed into the gray-scale waveforms according to EIA-608 Line 21 specifications. *See* Segment E of Figure 2. In one preferred embodiment, for each frame, the Field 1 packet waveform is created on both lines 1 and 482 of a picture image, while the Field 2 packet waveform is placed on lines 2 and 483. The rest of the picture image corresponding to each frame is set to black (zero gray-scale value).

In some embodiments, in order to create Line 21 waveforms that a non-linear editing system will properly convert into data readable by a closed-caption decoder or other decoder, the software must replicate the horizontal spacing of the Line 21 waveforms as it is digitized by the particular non-linear editing system that will be used to merge the video and control data. The horizontal spacing of the Line 21 waveform has been found to differ among computer video editing system vendors but appears to be the same within a specific vendor's hardware line. For example, the mid-point of the maximum value of first Clock Run-In cycle is located at pixel 35 for an Avid video NLE system with the Avid Broadcast Video Board (ABVB) hardware board set. The Apple Macintosh based Media 100 video NLE system with software Release 5 places the midpoint of the maximum of the first Clock Run-In cycle at pixel 26. The pixel-column positions associated with the Clock Run-In, Start Bits and Character Bits can be determined for any non-linear editing system by digitizing a known reference signal containing closed caption data in a digital non-linear editing system and examining the picture line corresponding to Line 21 of the digitized signal. An inverse computer software algorithm can then be created to replicate the horizontal spacing of the Clock Run-In, Start Bits and Character Bits for that particular editing system.

Similar to the horizontal spacing of the gray-scale waveform discussed above, the gray-scale data values of the waveform, in some embodiments, must be configured to achieve the appropriate luminance levels (measured in IRE units). It is important to determine the gray-scale data value that causes the NLE system output to have a video level of 50 IREs, according to EIA-608 specifications. Of course, different specifications may require different IRE levels. This gray-scale data value is the gray scale value for the pictorial

representation of the waveform. For example, when using an AVID system, the gray-scale value of 30,000 for 16-bit gray scale achieves a 50 IRE level. The minimum gray scale value is set to zero. The gray-scale data values along the pixel points in the waveform vary between this maximum and minimum according to the EIA-608 specification. *See* Figure 2. More specifically, the gray scale image is created by first determining the relative intensity of the pixel that is needed at a given location to create the desired EIA-608 specification IRE level and then by inserting that data value into the red, blue, and green components of the pixel.

To allow for different techniques for merging the images containing the pixel representations of the control data and the underlying video data, in one embodiment, the resulting pixel representations of the control data in the Field 1 packet array are copied onto line 0 and line 482 of the output image. The binary representation of the two character data values from the Field 2 array are encoded in a similar manner and put onto line 1 and line 483. However, different NLEs from the same or other manufactures may require that the Field 1 and Field 2 packet waveforms be placed on different rows of the picture image, because the hardware mapping from a row of the picture image to Line 21 of the VBI varies among NLE manufacturers and NLE systems. For example, it has been found that, when using an NLE system including the Avid Media Composer with an ABVB set, the Line 21 waveforms must be mapped on lines 0 and 482 for the Field 1 waveform and lines 1 and 483 for the Field 2 waveform. A non-linear editing system that includes the Media 100 release 5 software for the Macintosh platform requires that the Field 1 waveform be placed on lines 1 and 481 and that the Field 2 waveform be placed on lines 2 and 482. In addition, the duplication of the waveform on different lines in the picture image is not essential to the practice of the present invention. Rather, as described below, the placement of duplicate waveforms allows for use of different editing techniques to merge the video data file with the control data image file.

As discussed above, when output is requested the apparatus fills the two linear data arrays and then generates either multimedia or movie file or sequentially numbered image files from these arrays. This resulting video image file(s) is(are) then imported into a video non-linear editing system and placed on that system's editing time line. The data contained

within the imported file is then combined with the underlying video material using the video NLE system. In one preferred form, combining the imported data with video program material within the video NLE system can be accomplished with a Picture-In-Picture effect or with an Alpha channel image. Of course, other techniques such as image cropping could be used.

The PIP effect is used to position the Line 21 data value line contained within the imported control data image onto an image line within the video NLE system that will be mapped onto Line 21 of the VBI on output. For the Avid video NLE, the PIP effect is used to position data lines 482 and 483 of the control data image to the top of the underlying video image. The remainder of the control data image physically above lines 482 and 483 in the PIP lies beyond the mappable regions of the non-linear editing system and, thus, do not merge in the resulting video. On a Macintosh based Avid Media Composer NLE system running Version 7 Avid Software, the Size Parameter of the PIP should be set to 100 percent and the Vertical position should be set to -991. In one embodiment, the resulting merged video data file is output to a video playback and recording device 40, such as a VCR, which records the encoded video data.

An alternate method of combining the imported control data with video material is through the use of an alpha mask image. Typically, the pixel positions in the alpha mask that are white, a 255 luminance value, allow the imported image to be output and pixel positions that are black, a 0 luminance value, allow the program video material to be output. For this operation on an AVID NLE system, for example, the gray scale lines that are located on row 0 and row 1 in the picture image of the imported file are used by the NLE system. The alpha channel mask is configured to select these two rows.

C. Digital Video (DV) formats

Embodiments of the present invention also encode control data into digital video file formats, such as video data stored on helical-scan digital video cassettes. Unlike analog video data streams, digital video data streams do not include a vertical blanking interval. Rather, the digital video data stream includes video, audio, and auxiliary data portions for each frame in repeating succession. Accordingly, the digital video player constructs each

frame including the vertical blanking interval as part of the streaming analog signal, which is output during playback of the digital video data stream.

According to the invention, binary representations of control data as contained in the data structure, rather than pixel representations described above, are inserted into the auxiliary portions of the digital video data stream. In one embodiment, control data is inputted into the system. In one embodiment, the control data is converted into the appropriate byte code representations and temporally arranged in a data structure comprising Field 1 and Field 2 packet data arrays, as described above. The control data information within the data structure is then merged with the underlying digital video data file.

One embodiment of the present invention employs the data structures and methods of inserting control data into such data structures, described above. (See Figures 3A, 3B, and 3C, as well as related description). Unlike the embodiment described above. however, computer 20 does not include data file waveform generator 26 that generates pixel representations of control data. Rather, computer 20 includes a module that inserts control data from the data structure into the digital video data stream according to the standards and protocols required by the digital video format being used. For example and in one embodiment, control data is merged according to the protocols specified in International Standard International Electrotechnical Commission (IEC) 61834, parts 1-5, of Helical-Scan Digital Video Cassette Recording System Using 6,35 mm Magnetic Tape for Consumer Use (1st Ed. 1998), which is incorporated herein by reference, and is available at "http://www.iec.ch/."

Figure 17 illustrates an embodiment of a method for inserting control data into a digital video data stream. In one embodiment, digital video data is loaded and processed on a frame-by-frame basis in computer 20 where the control data is inserted into the copy of the data file corresponding to a frame stored in RAM. The data file corresponding to the frame can then be output to a computer readable medium such as a digital video cassette or stored in a digital video data file resident on the hard drive of computer 20. In one embodiment, the pointers for the Field 1 and Field 2 packet arrays are initialized (step 302). If the pointer is not at the end of the Field 1 and Field 2 packet arrays (step 304), a data file

module retrieves the control data from the Field 1 and/or Field 2 packet arrays (step 306) and inserts the control data into the auxiliary data portion of the digital video data stream at the appropriate frame or time point (step 308). In one embodiment, the digital video data file is further edited to ensure that the control data is compatibly merged (step 310), such as changing checksum values to reflect the addition of control data. The changes required to the digital video data file depend on the digital video file format being used. For example, the International Standards, discussed above, describes the changes required to add data to the auxiliary data portion of the digital video data file. This process is repeated until the pointer is at the end of the data arrays. (See Figure 17, steps 312, 304, and 322.)

Although embodiments have been described as implemented on an Apple Macintosh computer, porting the software algorithms to other computer architectures, such as the Intel based PC architecture, and using PC data and extensions to the AVI file format similar to QuickTime, for example, is straightforward. Moreover, the present invention is not limited to the file formats described above. Use of other file formats, such as JPEG, TARGA, TIFF, among others, are also capable of being implemented in the present invention. Although the use of only two video non-linear editing systems, one from Avid Technology and one from Media 100, have been described, the present invention, is easily adapted to any hardware device that can output an array of data storage elements to a line or lines of the vertical blanking interval.